State Estimation for Robotics Solution

Thomas
My personal solution of exercises in "State Estimation for Robotics", D. Barfoot.
Post on https://github.com/taohu1994/State_Estimation_for_Robotics_solution
May 2021

1 Chapter 2. Primer on Probability Theory

1.1 2.5.1

Let $\mathbf{u} = [u_1, \ldots, u_n]^T$, $\mathbf{v} = [v_1, \ldots, v_n]^T$, we have

$$\mathbf{u}^T\mathbf{v} = u_1v_1 + u_2v_2 + \cdot + u_nv_n$$

and

$$\mathbf{v}^T\mathbf{u} = \begin{bmatrix} v_1u_1 & \ldots & \ldots & v_1u_n \\ \vdots & v_2u_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ v_nu_1 & \ldots & \ldots & v_nu_n \end{bmatrix}.$$

It is clear that $\mathbf{u}^T\mathbf{v} = \mathrm{tr}(\mathbf{v}^T\mathbf{u})$.

1.2 2.5.2

Recapping the equation (2.29),

$$I(\mathbf{x}, \mathbf{y}) = H(\mathbf{x}) + H(\mathbf{y}) - H(\mathbf{x}, \mathbf{y}).$$

For two independent variables $\mathbf{x}, \mathbf{y}$, we have

$$I(\mathbf{x}, \mathbf{y}) = 0.$$

Then, it is clear that

$$H(\mathbf{x}) + H(\mathbf{y}) = H(\mathbf{x}, \mathbf{y}).$$

1.3 2.5.3

For a Gaussian random variable, $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, Its covariance matrix could be presented

$$\begin{aligned} \boldsymbol{\Sigma} &= E\{(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T\} \\ &= E\{\mathbf{x}\mathbf{x}^T\} - 2E\{(\mathbf{x}\boldsymbol{\mu}^T\} + \boldsymbol{\mu}\boldsymbol{\mu}^T \\ &= E\{\mathbf{x}\mathbf{x}^T\} - \boldsymbol{\mu}\boldsymbol{\mu}^T. \end{aligned}$$

We rewrite it as

$$E\{\mathbf{x}\mathbf{x}^T\} = \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^T.$$

1.4 2.5.4

We have

$$
\begin{aligned}
E[\mathbf{x}] &= \int_{-\infty}^{+\infty} \mathbf{x}(2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}} \exp\big(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\big)d\mathbf{x} \\
&= \int_{-\infty}^{+\infty} (\mathbf{y}+\boldsymbol{\mu})(2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}} \exp\big(-\frac{1}{2}\mathbf{y}^T\mathbf{\Sigma}^{-1}\mathbf{y}\big)d\mathbf{y} \\
&= \int_{-\infty}^{+\infty} \mathbf{y}(2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}} \exp\big(-\frac{1}{2}\mathbf{y}^T\mathbf{\Sigma}^{-1}\mathbf{y}\big)d\mathbf{y} + \int_{-\infty}^{+\infty} \boldsymbol{\mu}(2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}} \exp\big(-\frac{1}{2}\mathbf{y}^T\mathbf{\Sigma}^{-1}\mathbf{y}\big)d\mathbf{y},
\end{aligned}
$$

where we use the substitution $\mathbf{y} = \mathbf{x} - \boldsymbol{\mu}$. The first integral part could be rewritten

$$
\begin{aligned}
&\int_{-\infty}^{+\infty} \mathbf{y}(2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}} \exp\big(-\frac{1}{2}\mathbf{y}^T\mathbf{\Sigma}^{-1}\mathbf{y}\big)d\mathbf{y} \\
&= \int_{-\infty}^{0} \mathbf{y}(2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}} \exp\big(-\frac{1}{2}\mathbf{y}^T\mathbf{\Sigma}^{-1}\mathbf{y}\big)d\mathbf{y} + \int_{0}^{+\infty} \mathbf{y}(2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}} \exp\big(-\frac{1}{2}\mathbf{y}^T\mathbf{\Sigma}^{-1}\mathbf{y}\big)d\mathbf{y} \\
&= -\int_{0}^{+\infty} -\mathbf{y}(2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}} \exp\big(-\frac{1}{2}\mathbf{y}^T\mathbf{\Sigma}^{-1}\mathbf{y}\big)d(-\mathbf{y}) + \int_{0}^{+\infty} \mathbf{y}(2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}} \exp\big(-\frac{1}{2}\mathbf{y}^T\mathbf{\Sigma}^{-1}\mathbf{y}\big)d\mathbf{y} \\
&= 0.
\end{aligned}
$$

Therefore, considering the substitution $\mathbf{z} = \sqrt{\frac{1}{2}}\mathbf{\Sigma}^{-\frac{1}{2}}\mathbf{x}$, the mean became

$$
\begin{aligned}
E[\mathbf{x}] &= \boldsymbol{\mu}(2\pi)^{-\frac{N}{2}}2^{\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}}|\mathbf{\Sigma}|^{\frac{1}{2}} \int_{-\infty}^{+\infty} \exp\big(-\mathbf{z}^T\mathbf{z}\big)d\mathbf{z} \\
&= \boldsymbol{\mu}\pi^{-\frac{N}{2}} \int_{-\infty}^{+\infty} \exp\big(-\mathbf{z}^T\mathbf{z}\big)d\mathbf{z} \\
&= \boldsymbol{\mu}.
\end{aligned}
$$

Note that $\int_{-\infty}^{+\infty}\exp(-x^2)dx = \sqrt{\pi}$.

1.5 2.5.5

We present the variance as follow:

$$
\begin{aligned}
Var[\mathbf{x}] &= (2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}} \int_{-\infty}^{+\infty} (\mathbf{x}-\boldsymbol{\mu})^T(\mathbf{x}-\boldsymbol{\mu}) \exp\big(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\big)d\mathbf{x} \\
&= (2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}}|\mathbf{\Sigma}|^{\frac{1}{2}}2^{\frac{N}{2}}2^N\mathbf{\Sigma} \int_{-\infty}^{+\infty} \mathbf{y}^T\mathbf{y}\exp\big(-\mathbf{y}^T\mathbf{y}\big)d\mathbf{y} \\
&= (\pi)^{-\frac{N}{2}}\mathbf{\Sigma}2^N 2\int_{0}^{+\infty} \frac{1}{2}\mathbf{t}^{\frac{1}{2}}\exp\big(-\mathbf{t}\big)d\mathbf{t} \\
&= \pi^{-\frac{N}{2}}\Gamma(\frac{3}{2})^N 2^N\mathbf{\Sigma} \\
&= \pi^{-\frac{N}{2}}(\frac{\sqrt{\pi}}{2})^N 2^N\mathbf{\Sigma} \\
&= \mathbf{\Sigma}.
\end{aligned}
$$

Here we use the substitutions $\mathbf{y} = \sqrt{\frac{1}{2}}\boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{x} - \boldsymbol{\mu})$ and $\mathbf{t} = \mathbf{y}^T\mathbf{y}$. Note that

$$\int_0^{+\infty} \frac{1}{2}\mathbf{t}^{\frac{1}{2}} \exp\big(-\mathbf{t}\big)d\mathbf{t} = \Gamma(\frac{3}{2}),$$

where $\Gamma(\cdot)$ is the Gamma function.

1.6 2.5.6

Product of K Gaussian PDFs could be written

$$\begin{aligned}
&\eta\prod_{k=1}^{K} \exp\Big(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_k)\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_k)^T\Big)\\
&= \eta \exp\big(-\frac{1}{2}\sum_{k=1}^{K}(\mathbf{x}-\boldsymbol{\mu}_k)\boldsymbol{\Sigma}_k^{-1}(\mathbf{x}-\boldsymbol{\mu}_k)^T\big)\\
&= \eta \exp\Big(-\frac{1}{2}\big(\mathbf{x}^T\sum_{i=1}^{K}\boldsymbol{\Sigma}_k^{-1}\mathbf{x} - \mathbf{x}^T\sum_{k=1}^{K}\boldsymbol{\Sigma}_k^{-1}\boldsymbol{\mu}_k - (\sum_{k=1}^{K}\boldsymbol{\mu}_k^T\boldsymbol{\Sigma}_k^{-1})\mathbf{x} + \sum_{k=1}^{K}\boldsymbol{\mu}_k^T\boldsymbol{\Sigma}_k^{-1}\boldsymbol{\mu}_k)\Big)\\
&= \eta \exp\Big(-\frac{1}{2}\big(\mathbf{x}^T\boldsymbol{\Sigma}^{-1}\mathbf{x} - \mathbf{x}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} - \boldsymbol{\mu}^T\boldsymbol{\Sigma}^{-1}\mathbf{x} + \boldsymbol{\mu}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\big)\Big)\exp\big(-\frac{1}{2}(\sum_{k=1}^{K}\boldsymbol{\mu}_k^T\boldsymbol{\Sigma}_k^{-1}\boldsymbol{\mu}_k - \boldsymbol{\mu}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu})\big)\\
&= \eta' \exp\big(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\big),
\end{aligned}$$

where η' is a normalization constant to enforce the axiom of the total probability, $\boldsymbol{\Sigma}^{-1} = \sum_{k=1}^{K}\boldsymbol{\Sigma}_k^{-1}$ and $\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} = \sum_{k=1}^{K}\boldsymbol{\Sigma}_k^{-1}\boldsymbol{\mu}_k$.

1.7 2.5.7

Let $\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i$ denote the mean and variance of variable $\mathbf{x}_i$. As $\forall i,j \in [1,K], i \neq j, \mathbf{x}_i, \mathbf{x}_j$ are statistically independent. We have $E[(\mathbf{x}_i - \boldsymbol{\mu}_i)(\mathbf{x}_i - \boldsymbol{\mu}_i)^T] = \boldsymbol{\Sigma}_i$ and $E[(\mathbf{x}_i - \boldsymbol{\mu}_i)(\mathbf{x}_j - \boldsymbol{\mu}_j)^T] = \mathbf{0}$. The mean of $\mathbf{x}$ could be presented

$$E[\mathbf{x}] = \sum_{(i=1)}^{K}\omega_i E[\mathbf{x}_i] = \sum_{(i=1)}^{K}\omega_i\boldsymbol{\mu}_i = \boldsymbol{\mu}$$

The variance of $\mathbf{x}$ could be presented

$$\begin{aligned}
Var(\mathbf{x}) &= E[\big(\omega_1(\mathbf{x}_1-\boldsymbol{\mu}_1)+\cdots+\omega_K(\mathbf{x}_K-\boldsymbol{\mu}_K)\big)\big(\omega_1(\mathbf{x}_1-\boldsymbol{\mu}_1)+\cdots+\omega_K(\mathbf{x}_K-\boldsymbol{\mu}_K)\big)^T]\\
&= \sum_{i=1}^{K}\omega_i^2 E\{(\mathbf{x}_i-\boldsymbol{\mu}_i)(\mathbf{x}_i-\boldsymbol{\mu}_i)^T\} + \sum_{\forall i,j\in[1,K], i\neq j}\omega_i\omega_j E\{(\mathbf{x}_i-\boldsymbol{\mu}_i)(\mathbf{x}_j-\boldsymbol{\mu}_j)^T\}\\
&= \sum_{i=1}^{K}\omega_i^2 E\{(\mathbf{x}_i-\boldsymbol{\mu}_i)(\mathbf{x}_i-\boldsymbol{\mu}_i)^T\}\\
&= \sum_{i=1}^{K}\omega_i^2\boldsymbol{\Sigma}_i.
\end{aligned}$$

1.8 2.5.8

Note $\mathbf{x} = [x_1, x_2, \ldots, x_K]^T$. For scalar $x_i \in \mathbf{x}$, we have $x_i \sim \mathcal{N}(0, 1)$. And $\forall i, j \in [1, K], i \neq j, E\{x_i x_i\} = 1, E\{x_i x_j\} = 0$. We write the mean of $\mathbf{y}$ as

$$\begin{aligned} E\{\mathbf{x}^T\mathbf{x}\} &= E\{x_1 x_1 + \cdots + x_K x_K\} \\ &= E\{x_1 x_1\} + \cdots + E\{x_K x_K\} \\ &= \sum_{i=1}^{K} E\{x_i x_i\} \\ &= K. \end{aligned}$$

The variance of $\mathbf{y}$,

$$\begin{aligned} \mathrm{Var}(\mathbf{y}) &= E\{(\mathbf{x}^T\mathbf{x} - K)(\mathbf{x}^T\mathbf{x} - K)^T\} \\ &= E\{\mathbf{x}^T\mathbf{x}\mathbf{x}^T\mathbf{x}\} - 2E\{\mathbf{x}\mathbf{x}^T\}K + K^2 \\ &= E\{(x_1 x_1 + \cdots + x_K x_K)(x_1 x_1 + \cdots + x_K x_K)\} - K^2 \\ &= E\{\sum_{i=1}^{K} x_i x_i x_i x_i\} + E\{\sum_{\forall i, j \in [1,K], i \neq j} x_i x_i x_j x_j\} - K^2 \\ &= KE\{x_i x_i x_i x_i\} + (K^2 - K)E\{x_i x_i x_j x_j\} - K^2 \\ &= 2K. \end{aligned}$$

Note that here we use the Isserlis' Theorem and equation (2.40), that

$$E\{x_i x_i x_i x_i\} = 3E\{x_i x_i\}E\{x_i x_i\} = 3.$$

and

$$E\{x_i x_i x_j x_j\} = E\{x_i x_i\}E\{x_j x_j\} + 2E\{x_i x_j\}E\{x_i x_j\} = 1.$$

1.9 ex 2.4

The integrate of an odd function is zero in the symmetric interval.

$$
\begin{aligned}
E(\mathbf{x}) &= \int_{-\infty}^{+\infty} \mathbf{x} p(\mathbf{x})\, \mathrm{d}\mathbf{x} \\
&= \int_{-\infty}^{+\infty} \frac{\mathbf{x}}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right) \mathrm{d}\mathbf{x} \\
&= \int_{-\infty}^{+\infty} \frac{\mathbf{y}+\boldsymbol{\mu}}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2}\mathbf{y}^T \boldsymbol{\Sigma}^{-1}\mathbf{y}\right) \mathrm{d}\mathbf{y} \\
&= \underbrace{\int_{-\infty}^{+\infty} \frac{\mathbf{y}}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2}\mathbf{y}^T \boldsymbol{\Sigma}^{-1}\mathbf{y}\right) \mathrm{d}\mathbf{y}}_{0} + \int_{-\infty}^{+\infty} \frac{\boldsymbol{\mu}}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2}\mathbf{x}^T \boldsymbol{\Sigma}^{-1}\mathbf{x}\right) \mathrm{d}\mathbf{x} \\
&= \boldsymbol{\mu} \underbrace{\int_{-\infty}^{+\infty} \frac{1}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2}\mathbf{y}^T \boldsymbol{\Sigma}^{-1}\mathbf{y}\right) \mathrm{d}\mathbf{y}}_{\mathbf{y} \sim \boldsymbol{N}(0, \boldsymbol{\Sigma})} \\
&= \boldsymbol{\mu} \underbrace{\int_{-\infty}^{+\infty} p(\mathbf{y})\, \mathrm{d}\mathbf{y}}_{1} \\
&= \boldsymbol{\mu}
\end{aligned}
\tag{1}
$$

1.10 ex 2.5

$$
\begin{aligned}
E\left((\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^T\right) &= \int_{-\infty}^{\infty} (\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^T \frac{1}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}) \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})^T\right) \mathrm{d}\mathbf{x} \\
&= \frac{1}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}}} \underbrace{\int_{-\infty}^{\infty} d\left(-\boldsymbol{\Sigma}(\mathbf{x}-\boldsymbol{\mu}) \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}) \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})^T\right)\right)}_{\text{odd},=0} \\
&\quad + \frac{1}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}}} \int_{-\infty}^{\infty} \boldsymbol{\Sigma} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}) \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})^T\right) \mathrm{d}\mathbf{x} \\
&= \mathbf{0} + \boldsymbol{\Sigma} = \boldsymbol{\Sigma}
\end{aligned}
\tag{2}
$$

$$
\begin{aligned}
\mathrm{Cov}(\mathbf{x}) &= E((\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^T) \\
&= E\left(\sum_{k=1}^{K}(w_k\mathbf{x}_k - w_k\boldsymbol{\mu}_k)\sum_{k=1}^{K}(w_k\mathbf{x}_k - w_k\boldsymbol{\mu}_k)^T\right) \\
&= E\left(\sum_{k=1}^{K} w_k^2\left(\mathbf{x}_k-\boldsymbol{\mu}_k\right)\left(\mathbf{x}_k-\boldsymbol{\mu}_k\right)^T + \sum_{m=1,n=1,m\neq n}^{K} w_m w_n\left(\mathbf{x}_m-\boldsymbol{\mu}_m\right)\left(\mathbf{x}_n-\boldsymbol{\mu}_n\right)^T\right) \\
&= \sum_{k=1}^{K} w_k^2 \mathrm{Cov}\left(\mathbf{x}_k\right) + \underbrace{E\left(\sum_{m=1,n=1,m\neq n}^{K} w_m w_n\left(\mathbf{x}_m-\boldsymbol{\mu}_m\right)\left(\mathbf{x}_n-\boldsymbol{\mu}_n\right)^T\right)}_{\text{independent,=0}} \\
&= \sum_{k=1}^{K} w_k^2 \mathrm{Cov}\left(\mathbf{x}_k\right)
\end{aligned} \tag{3}
$$

References